CHECK

MST281 10

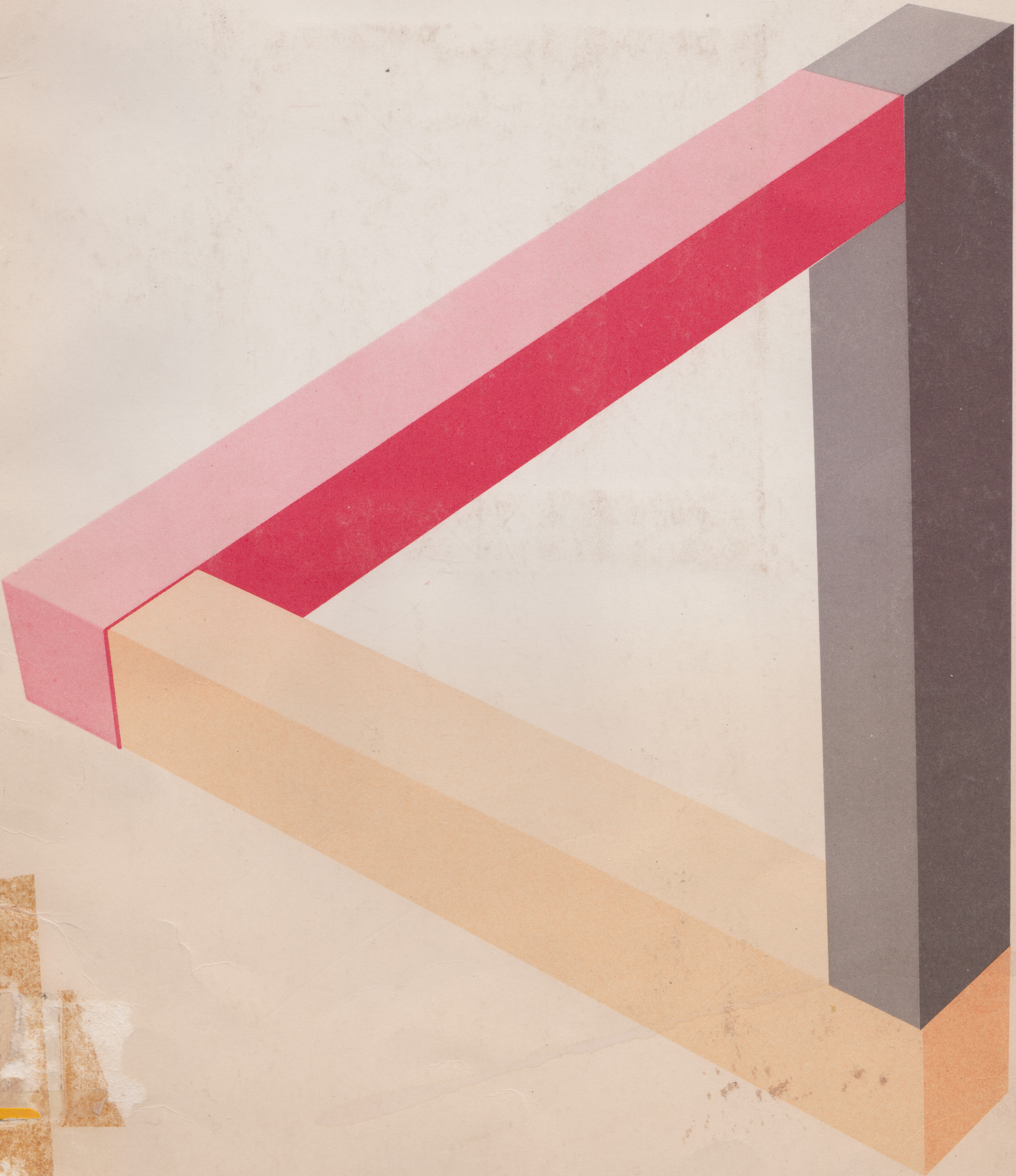
THE OPEN UNIVERSITY

Mathematics/Science/Technology

An Inter-Faculty Second Level Course

Elementary Mathematics for Science and Technology Unit 10

Taylor Approximation



The Open University

Mathematics/Science/Technology
An Inter-Faculty Second Level Course

Elementary Mathematics for Science and Technology

Unit 10 TAYLOR APPROXIMATION

Prepared by the Course Team

The Open University Press

The Open University Press Walton Hall Bletchley Bucks

First published 1972
Copyright © 1972 The Open University

All rights reserved. No part of this work may
be reproduced in any form, by mimeograph
or any other means, without permission in
writing from the publishers.

Designed by the Media Development Group of the Open University.

Printed in Great Britain by
Martin Cadbury Printing Group

SBN 335 01215 9

This text forms part of the correspondence element of an Open University Second Level Course. The complete list of units in the course is given at the end of this text.

For general availability of supporting material referred to in this text, please write to the Director of Marketing, The Open University, Walton Hall, Bletchley, Buckinghamshire.

Further information on Open University courses may be obtained from the Admissions Office, The Open University, P.O. Box 48, Bletchley, Buckinghamshire.

1.1

Contents

Page

10.0 Introduction 4

10.1 Core Material 5

10.1.1 The Tangent Approximation 5
10.1.2 The Quadratic Taylor Approximation 9
10.1.3 The General Taylor Approximation 14
10.1.4 Error in the Tangent Approximation 16

10.2 Optional Material 19

10.2.1 The Convergence of an Iterative Method 19
10.2.2 The Newton-Raphson Process 21
10.2.3 Power Series 23
10.2.4 Taylor's Theorem 26
10.2.5 Maxima and Minima of Functions of Two Real Variables 27

10.3 Summary 30

10.0 INTRODUCTION

This unit is concerned with the approximation of certain "well-behaved" functions of one real variable by polynomial functions which exhibit a particular pattern. It discusses the question: given a function f and its image and derivatives at a point a in its domain, what can we say about the image $f(x)$, where x is "near to" a?

The unit introduces the general Taylor approximation of degree n for a function f which possesses derived functions, $f', f'', f^{(3)}, \ldots, f^{(n)}$. This approximation expresses $f(x)$ as a polynomial of degree n in $(x - a)$, with coefficients which involve only real numbers and the derivatives of f at a.

The *core material* begins by discussing the linear Taylor approximation (i.e. the tangent approximation) and the quadratic Taylor approximation. The pattern which emerges leads us to consider the general Taylor approximation of degree n. We find that, knowing $f(a), f'(a), \ldots, f^{(n)}(a)$ $f^{(n+1)}$, we *can* say something useful about $f(x)$, where x is "near" a. This leads us to discuss what we mean by "near" i.e. the validity of the Taylor approximation. In section 10.1.4 we discuss the error in the tangent approximation and give an expression for the error bound, and indicate the error analysis for the general case. We find that, in many cases, to stay within a particular error bound we have to place a restriction on x.

The *optional material* falls into three main parts.

In sections 10.2.1 and 10.2.2 we take up the discussion of iterative methods of solving equations, and we answer some of the questions raised in our discussion of iterative methods in *Units 1* and *2*. In section 10.2.1 we consider sufficient conditions for the convergence of an iterative method. In section 10.2.2 we consider a *particular* iterative method, called the *Newton-Raphson* method.

In section 10.2.3 we extend our earlier work on the convergence of infinite series and consider *power series*, that is, series of the form

$$a_0 + a_1(x - a) + a_2(x - a)^2 + a_3(x - a)^3 + \cdots$$

The convergence of such a series usually depends on the value assigned to the variable x. In section 10.2.4 we consider Taylor's series and its convergence.

Finally, in section 10.2.5, we return to the problem of the location and determination of the local extrema of functions of two real variables. We use Taylor's series (for one real variable) to derive a test for the determination of the nature of these extrema. (See *Unit 7*, section 7.1.7, Note c(4).)

10.1 CORE MATERIAL

10.1.1 The Tangent Approximation

(a) Purpose

To explain how a differentiable function f of one real variable can be approximated in the neighbourhood of a point a in its domain by a linear function whose graph is the tangent to the graph of f at a.

(b) Set Book

Read section **2.5.0** (p.**115**). Study section **2.5.1** (pp.**115-119**) and complete Exercises 1 and 2.

(c) Notes

(1) We met the geometrical interpretation of the derivative of f at a as the tangent to the graph of f at a in section **1.8.2** (**pp.177-178**).

(2) Let t be the tangent approximation to f about $x = a$; i.e.

$$t : x \longmapsto f(a) + (x - a)f'(a).$$

Then the derived function of t is

$$t' : x \longmapsto f'(a).$$

Hence

$$t(a) = f(a) \qquad \text{(i)}$$

and

$$t'(a) = f'(a). \qquad \text{(ii)}$$

Conversely, let t be a linear function,

$$t : x \longmapsto b_0 + b_1 x \qquad (x \in R).$$

Then t is the tangent approximation to f at a if the graph of t passes through the point $(a, f(a))$, i.e.

$$t(a) = b_0 + b_1 a = f(a), \qquad \text{(i)}$$

and the slopes of the graphs of f and t at a are the same, i.e.

$$t'(a) = b_1 = f'(a). \qquad \text{(ii)}$$

From equation (i), we have

$$b_0 = f(a) - b_1 a;$$

hence, from equation (ii) it follows that t is the function

$$t : x \longmapsto f(a) + f'(a)(x - a) \qquad (x \in R).$$

This result is in agreement with Equation (1) on p.**118**.

(The importance of this approach stems from the fact that it can be generalized to derive higher order polynomial approximations. We shall consider these in sections 10.1.2 and 10.1.3.)

We have seen that equations (i) and (ii) completely determine the tangent approximation to f about $x = a$, and conversely.

(3) If t is the tangent approximation to f about $x = a$, then

$$\lim_{x \to a} \big(f(x) - t(x)\big) = f(a) - t(a) = 0,$$

by equation (i).

Also,

$$\begin{aligned}\lim_{x\to a}\frac{f(x)-t(x)}{x-a} &= \lim_{x\to a}\frac{f(x)-f(a)-(x-a)f'(a)}{x-a}\\ &= \lim_{x\to a}\frac{f(x)-f(a)}{x-a}-f'(a)\\ &= f'(a)-f'(a)\\ &= 0.\end{aligned}$$

We now have the two results:

$$\lim_{x\to a}\big(f(x)-t(x)\big)=0 \tag{iii}$$

and

$$\lim_{x\to a}\frac{f(x)-t(x)}{x-a}=0. \tag{iv}$$

Conversely, if t is any linear function which satisfies equations (iii) and (iv), then, from equation (iii), it follows that

$$f(a)=t(a),$$

since f and t are continuous at a.

We may write equation (iv) in the form

$$\lim_{x\to a}\frac{f(x)-f(a)-\big(t(x)-t(a)\big)}{x-a}=0,$$

i.e.

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a}+\lim_{x\to a}\frac{t(x)-t(a)}{x-a}=0;$$

hence

$$f'(a)=t'(a).$$

Hence t is the tangent approximation to f about $x=a$. (See Note (2)).

Consequently, t is the tangent approximation to f about $x=a$ if and only if equations (iii) and (iv) are satisfied. (Again, these results can be generalized for higher order polynomial approximations.)

(4) Equation (iv) gives us an upper bound on the difference $|f(x)-t(x)|$ between the actual value of the function f at x and the approximate value given by $t(x)$, when x is "near" a. Because $\lim_{x\to a}\frac{f(x)-t(x)}{x-a}=0$, then, for x sufficiently near a,

$$\frac{|f(x)-t(x)|}{|x-a|}<1.$$

(otherwise the limit could not be zero). Hence

$$|f(x)-t(x)|<|x-a|, \tag{v}$$

for x sufficiently near a.

If we define the error in the approximation at x to be the difference $|f(x)-t(x)|$, then inequality (v) gives an upper bound on the error, provided x is near a. In the special case when we consider the tangent approximation to f about $x=0$, inequality (v) becomes

$$|f(x)-t(x)|<|x|.$$

Of course, the phrase "x sufficiently near a" is very vague, but errors in approximations will be considered more rigorously in section **2.5.6**.

(d) Self-Assessment Questions

1 Derive an expression for the error in approximating the function

$$f : x \longmapsto x^3 \qquad (x \in R^+)$$

over an interval $[a - \delta, a + \delta] \subseteq R^+$ by the tangent approximation at the point $(a, f(a))$.

2 Derive an expression for the error in approximating the function

$$f : x \longmapsto \frac{1}{x} \qquad (x \in R^+)$$

over an interval $[a - \delta, a + \delta] \subseteq R^+$ by the tangent approximation at the point $(a, f(a))$.

3 Use your results for questions 1 and 2 to show that, in each case,

$$\text{error} \leqslant \tfrac{1}{2}|x - a|^2 \times \left(\begin{matrix}\text{maximum value of} \\ |f''(x)| \text{ in } [a - \delta, a + \delta]\end{matrix}\right).$$

Solutions

1 The derived function of f is

$$f' : x \longmapsto 3x^2 \qquad (x \in R^+),$$

so the tangent approximation at $(a, f(a))$ is

$$t : x \longmapsto a^3 + (x - a)3a^2 \qquad (x \in R^+).$$

The error in the approximation is

$$|f(x) - t(x)| = |x^3 - a^3 - (x - a)3a^2|.$$

If $x \in [a - \delta, a + \delta]$, then we can write $x = a + \varepsilon$, where $\varepsilon \in [-\delta, \delta]$. Thus

$$\begin{aligned}\text{error} &= |(a + \varepsilon)^3 - a^3 - \varepsilon 3a^2| \\ &= |3a\varepsilon^2 + \varepsilon^3| \\ &= |3a + \varepsilon|\varepsilon^2.\end{aligned}$$

2 The derived function of f is

$$f' : x \longmapsto -\frac{1}{x^2} \qquad (x \in R^+).$$

As in question 1, we can write $x = a + \varepsilon$, where $\varepsilon \in [-\delta, \delta]$. We have

$$\begin{aligned}\text{error} &= |f(x) - t(x)| \\ &= \left|\frac{1}{(a + \varepsilon)} - \frac{1}{a} + \frac{\varepsilon}{a^2}\right| \\ &= \left|\frac{\varepsilon^2}{(a + \varepsilon)a^2}\right| \\ &= \frac{\varepsilon^2}{a^2} \times \frac{1}{|a + \varepsilon|}.\end{aligned}$$

3 (i) In question 1,

$$f'' : x \longmapsto 6x \qquad (x \in R^+),$$

so

$$\tfrac{1}{2}|x - a|^2 \times \left(\begin{matrix}\text{maximum value of} \\ |f''(x)| \text{ in } [a - \delta, a + \delta]\end{matrix}\right)$$

is

$$\tfrac{1}{2}\varepsilon^2 \times 6(a + \delta) = (3a + 3\delta)\varepsilon^2.$$

From question 1,

$$\begin{aligned}\text{error} &= |3a+\varepsilon|\varepsilon^2\\ &\leqslant |3a+|\varepsilon||\varepsilon^2.\end{aligned}$$

Since $\delta > 0$, and $|\varepsilon| \leqslant \delta$, it follows that

$$\text{error} \leqslant \tfrac{1}{2}|x-a|^2 \times \left(\begin{matrix}\text{maximum value of}\\ |f''(x)| \text{ in } [a-\delta, a+\delta]\end{matrix}\right).$$

(ii) In question 2,

$$f'' : x \longmapsto \frac{2}{x^3} \qquad (x \in R^+),$$

so

$$\tfrac{1}{2}|x-a|^2 \times \left(\begin{matrix}\text{maximum value of}\\ |f''(x)| \text{ in } [a-\delta, a+\delta]\end{matrix}\right)$$

is

$$\frac{1}{2}\varepsilon^2 \times \frac{2}{(a-\delta)^3} = \frac{\varepsilon^2}{(a-\delta)^3}.$$

From question 2,

$$\begin{aligned}\text{error} &= \frac{\varepsilon^2}{a^2|a+\varepsilon|}\\ &\leqslant \frac{\varepsilon^2}{a^2(a-\delta)}.\end{aligned}$$

Since $\delta > 0$ and $a > 0$, it follows that $a^2(a-\delta) \geqslant (a-\delta)^3$, and therefore

$$\text{error} \leqslant \tfrac{1}{2}|x-a|^2 \times \left(\begin{matrix}\text{maximum value of}\\ |f''(x)| \text{ in } [a-\delta, a+\delta]\end{matrix}\right).$$

(e) Terminology

Defined in this section:

linear Taylor approximation (tangent approximation)

(f) Additional Exercise (OPTIONAL MATERIAL)

Use the tangent approximation to compute the following values:

(i) $(25.018)^{1/2}$
(ii) sin 29°

Solution

(i) We use the tangent approximation t to $f : x \longmapsto x^{1/2}$ about $x = 25$.

Then

$$t : x \longmapsto \sqrt{25} + (x-25)f'(25) \qquad (x \in R^+).$$

Now

$$f' : x \longmapsto \frac{1}{2x^{1/2}}$$

and so $f'(25) = 1/10$.

Hence

$$t : x \longmapsto 5 + \tfrac{1}{10}(x-25)$$

and

$$\begin{aligned}t(25.018) &= 5 + \tfrac{1}{10}(0.018)\\ &= 5 + 0.0018\\ &= 5.0018\end{aligned}$$

($(25.018)^{1/2}$ is 5.0017997 correct to 7 decimal places.)

(ii) $\sin 29° = \sin (30 - 1)° = \sin\left(\frac{\pi}{6} - \frac{\pi}{180}\right)$.

We use the tangent approximation t to $f : x \longmapsto \sin x$ about $x = \pi/6$. Then

$$t : x \longmapsto \sin\left(\frac{\pi}{6}\right) + \left(x - \frac{\pi}{6}\right)\cos\left(\frac{\pi}{6}\right) \qquad (x \in R)$$

i.e.

$$t : x \longmapsto \frac{1}{2} + \left(x - \frac{\pi}{6}\right)\frac{\sqrt{3}}{2}$$

whence

$$t\left(\frac{\pi}{6} - \frac{\pi}{180}\right) = \frac{1}{2} + \left(-\frac{\pi}{180}\right)\frac{\sqrt{3}}{2}$$

$$= 0.4849$$

(sin 29° is 0.4848 correct to 4 decimal places.)

10.1.2 The Quadratic Taylor Approximation

(a) Purpose

To explain how a twice differentiable function f of one real variable can be approximated in the neighborhood of a point a in its domain by a quadratic function q, dependent only upon $f(a)$, $f'(a)$ and $f''(a)$.

(b) Set Book

Study section **2.5.4** (**pp.122-124**) and complete Exercise 1. Ignore the first two sentences of this section for the moment. (The relation between the iterative method and the linear Taylor approximation is discussed in the Optional Material, section 10.2.1.)

(c) Notes

(1) Instead of regarding the quadratic Taylor approximation as the limit of a quadratic through three equally spaced points (as we have done on p.**123**) we can *start* by requiring that the quadratic function

$$q : x \longmapsto b_0 + b_1(x - a) + b_2(x - a)^2 \qquad (x \in R)$$

satisfies the conditions specified in Equations (1) (p.**123**). The derivation then follows that given, and is clearly a simple generalization of that given in section 10.1.1, Note c(2).

The equations

$$q(a) = f(a) \qquad \text{(i)}$$

$$q'(a) = f'(a) \qquad \text{(ii)}$$

$$q''(a) = f''(a) \qquad \text{(iii)}$$

completely determine the quadratic Taylor approximation to f about $x = a$, and conversely.

If you are interested in establishing the equivalence between the geometrical and algebraic interpretation of the quadratic Taylor approximation, then you can study the method described in section **2.7.7**, page **228**, to fit a quadratic to the points $(a, f(a))$, $((a + h, f(a + h))$ and $(a + 2h, f(a + 2h))$. The required quadratic is given by Equation (5), page **229**, where $x_k = a$, $x_{k+1} = a + h$ and $x_{k+2} = a + 2h$. Taking the limit as h approaches zero, and using the definitions of $f'(a)$ and $f''(a)$, it can be shown that Equation (5) on page **229** reduces to the quadratic Taylor approximation. This demonstrates the validity of the intuitive argument given on page **123**.

(2) If q is the quadratic approximation to f about $x = a$, then it can be shown that

$$\lim_{x \to a} f(x) - q(x) = 0 \qquad \text{(iv)}$$

$$\lim_{x \to a} \frac{f(x) - q(x)}{x - a} = 0 \qquad \text{(v)}$$

$$\lim_{x \to a} \frac{f(x) - q(x)}{(x - a)^2} = 0 \qquad \text{(vi)}$$

Conversely, if q is the quadratic function which satisfies equations (iv), (v) and (vi), then it can be shown that

$$q : x \longmapsto f(a) + (x - a)f'(a) + \tfrac{1}{2}(x - a)^2 f''(a).$$

Consequently q is the quadratic Taylor approximation to f about $x = a$ if and only if equations (iv), (v) and (vi) are satisfied.

(3) Equation (vi) gives us an upper bound for the error $|f(x) - q(x)|$ in the quadratic Taylor approximation.
Since

$$\lim_{x \to a} \frac{f(x) - q(x)}{(x - a)^2} = 0,$$

then, for any small positive number ε, it must be possible to find a positive number δ such that

$$\left| \frac{f(x) - q(x)}{(x - a)^2} \right| < \varepsilon$$

whenever $x \in [a - \delta, a + \delta]$.

Hence

$$|f(x) - q(x)| < \varepsilon |x - a|^2.$$

Taking $\varepsilon = 1$, we have

$$|f(x) - q(x)| < |x - a|^2,$$

for all x sufficiently "near" a.

(d) Self-Assessment Questions

1 Derive an expression for the error in approximating the function

$$f : x \longmapsto x^3 \qquad (x \in R^+)$$

over an interval $[a - \delta, a + \delta] \subseteq R^+$ by the quadratic Taylor approximation at the point $(a, f(a))$.

2 Derive an expression for the error in approximating the function

$$f : x \longmapsto \frac{1}{x} \qquad (x \in R^+)$$

over an interval $[a - \delta, a + \delta] \subseteq R^+$ by the quadratic Taylor approximation at the point $(a, f(a))$.

3 Use your results for questions 1 and 2 to show that, in each case,

$$\text{error} \leqslant \frac{1}{3!}|x - a|^3 \times \left(\begin{array}{c} \text{maximum value of} \\ |f^{(3)}(x)| \text{ in } [a - \delta, a + \delta] \end{array} \right).$$

4 Calculate an approximation to $(1.14)^{1/2}$ by

(i) the tangent approximation;
(ii) the quadratic Taylor approximation.

Compare your answers with the correct value, which is 1.067708 (correct to 6 decimal places).

Solutions

1 The first and second derived functions of f are

$$f' : x \longmapsto 3x^2 \qquad (x \in R^+),$$
$$f'' : x \longmapsto 6x \qquad (x \in R^+)$$

respectively.

The quadratic Taylor approximation at $(a, f(a))$ is

$$q : x \longmapsto a^3 + (x-a)3a^2 + \tfrac{1}{2}(x-a)^2 6a \qquad (x \in R^+).$$

The error in the approximation is

$$|f(x) - q(x)| = |x^3 - a^3 - (x-a)3a^2 - \tfrac{1}{2}(x-a)^2 6a|.$$

If $x \in [a-\delta, a+\delta]$, then we can write $x = a + \varepsilon$, where $\varepsilon \in [-\delta, \delta]$. The error then becomes

$$|(a+\varepsilon)^3 - a^3 - \varepsilon 3a^2 - \tfrac{1}{2}\varepsilon^2 6a| = |\varepsilon|^3.$$

2 The first and second derived functions of f are

$$f' : x \longmapsto -\frac{1}{x^2} \qquad (x \in R^+),$$
$$f'' : x \longmapsto \frac{2}{x^3} \qquad (x \in R^+)$$

respectively.

The quadratic Taylor approximation at $(a, f(a))$ is

$$q : x \longmapsto \frac{1}{a} - (x-a)\frac{1}{a^2} + \frac{1}{2}(x-a)^2\frac{2}{a^3} \qquad (x \in R^+).$$

The error in the approximation is

$$|f(x) - q(x)| = \left|\frac{1}{x} - \frac{1}{a} + (x-a)\frac{1}{a^2} - (x-a)^2\frac{1}{a^3}\right|.$$

If $x \in [a-\delta, a+\delta]$, then we can write $x = a + \varepsilon$, where $\varepsilon \in [-\delta, \delta]$. The error then becomes

$$\left|\frac{1}{a+\varepsilon} - \frac{1}{a} + \frac{\varepsilon}{a^2} - \frac{\varepsilon^2}{a^3}\right| = \left|\frac{\varepsilon}{a}\right|^3 \frac{1}{|a+\varepsilon|}.$$

3 (i) In question 1,

$$f^{(3)} : x \longmapsto 6 \qquad (x \in R^+),$$

so

$$\frac{1}{3!}|x-a|^3 \times \begin{pmatrix}\text{maximum value of} \\ |f^{(3)}(x)| \text{ in } [a-\delta, a+\delta]\end{pmatrix}$$

is

$$\tfrac{1}{6}|\varepsilon|^3 \times 6 = |\varepsilon|^3$$

Hence, in this case,

$$\text{error} = \frac{1}{3!}|x-a|^3 \times \begin{pmatrix}\text{maximum value of} \\ |f^{(3)}(x)| \text{ in } [a-\delta, a+\delta]\end{pmatrix}.$$

(ii) In question 2,

$$f^{(3)} : x \longmapsto -\frac{6}{x^4} \qquad (x \in R^+),$$

so

$$\frac{1}{3!}|x-a|^3 \times \begin{pmatrix}\text{maximum value of} \\ |f^{(3)}(x)| \text{ in } [a-\delta, a+\delta]\end{pmatrix}$$

is

$$\frac{1}{6}\varepsilon^3 \times \frac{6}{(a-\delta)^4} = \frac{\varepsilon^3}{(a-\delta)^4}.$$

From question 2, the error is

$$\left|\frac{\varepsilon}{a}\right|^3 \frac{1}{a+\varepsilon}.$$

Since $\delta > 0$ and $\varepsilon \geqslant -\delta$,

$$a^3(a+\varepsilon) > (a-\delta)^4,$$

hence

$$\text{error} \leqslant \frac{1}{3!}|x-a|^3 \times \left(\begin{matrix}\text{maximum value of}\\ |f^{(3)}(x)| \text{ in } [a-\delta, a+\delta]\end{matrix}\right).$$

4 If

$$f : x \longmapsto x^{1/2} \qquad (x \in R, x \neq 0),$$

then the tangent approximation about $x = 1$ is

$$t : x \longmapsto 1 + \tfrac{1}{2}(x-1) \qquad (x \in R, x \neq 0),$$

whence

$$\begin{aligned} t(1.14) &= 1 + \tfrac{1}{2}(0.14)\\ &= 1 + 0.07\\ &= 1.07 \end{aligned}$$

The quadratic Taylor approximation about $x = 1$ is

$$q : x \longmapsto f(1) + (x-1)f'(1) + \frac{(x-1)^2}{2}f''(1).$$

Now

$$f' : x \longmapsto \frac{1}{2x^{1/2}} \qquad (x \in R, x \neq 0)$$

$$f'' : x \longmapsto -\frac{1}{4x^{3/2}} \qquad (x \in R, x \neq 0)$$

and so

$$f''(1) = -\tfrac{1}{4},$$

giving

$$\begin{aligned} q : x &\longmapsto 1 + \tfrac{1}{2}(x-1) - \tfrac{1}{8}(x-1)^2,\\ q(1.14) &= 1 + \tfrac{1}{2}(0.14) - \tfrac{1}{8}(0.14)^2\\ &= 1 + 0.07 - 0.00245\\ &= 1.06755 \end{aligned}$$

We see that the quadratic Taylor approximation is more accurate for this value of x than the tangent approximation.

(e) Terminology

Defined in this section:

quadratic Taylor approximation

(f) Additional Exercises (OPTIONAL MATERIAL)

1 Use the quadratic Taylor approximation to calculate an approximation to cos (31°).

2 Find the quadratic Taylor approximation to the function

$$f : x \longmapsto 3x^2 - x + 6 \qquad (x \in R)$$

about the point $x = -1$.

What do you notice?

Solutions

1 $\cos(31°) = \cos\left(\frac{\pi}{6} + \frac{\pi}{180}\right).$

Let

$$f : x \longmapsto \cos x \qquad (x \in R).$$

Then

$$f' : x \longmapsto -\sin x \qquad (x \in R)$$
$$f'' : x \longmapsto -\cos x \qquad (x \in R)$$

and so the quadratic Taylor approximation about $x = \frac{\pi}{6}$ is

$$q : x \longmapsto f\left(\frac{\pi}{6}\right) + \left(x - \frac{\pi}{6}\right) f'\left(\frac{\pi}{6}\right) + \frac{1}{2}\left(x - \frac{\pi}{6}\right)^2 f''\left(\frac{\pi}{6}\right)$$

i.e.

$$q\left(\frac{\pi}{6} + \frac{\pi}{180}\right) = \frac{\sqrt{3}}{2} - \frac{\pi}{180} \times \frac{1}{2} - \left(\frac{\pi}{180}\right)^2 \frac{\sqrt{3}}{4}.$$

Thus

$$q\left(\frac{\pi}{6} + \frac{\pi}{180}\right) = 0.8572$$

(to four decimal places) and in fact this is the correct value to four decimal places.

2 We have

$$f : x \longmapsto 3x^2 - x + 6 \qquad (x \in R)$$
$$f' : x \longmapsto 6x - 1 \qquad (x \in R)$$
$$f'' : x \longmapsto 6 \qquad (x \in R)$$

so the quadratic Taylor approximation about $x = -1$ is

$$q : x \longmapsto f(-1) + (x+1)f'(-1) + \frac{(x+1)^2}{2} f''(-1)$$

i.e.

$$q : x \longmapsto 10 + (x+1)(-7) + \frac{(x+1)^2}{2} 6$$

i.e.

$$x \longmapsto 10 - 7x - 7 + 3x^2 + 6x + 3$$

i.e.

$$x \longmapsto 3x^2 - x + 6 \qquad (x \in R).$$

This is the same as the original function.

This should come as no surprise because both *f* and *q* are quadratics with

$$f(-1) = q(-1)$$
$$f'(-1) = q'(-1)$$
$$f''(-1) = q''(-1)$$

and as we observed earlier, equations (i)–(iii) in Note c(1) completely determine the quadratic approximation.

10.1.3 The General Taylor Approximation

(a) Purpose

To explain how an n-times differentiable function f of one real variable can be approximated in the neighbourhood of a point a in its domain by a polynomial function of degree n.

(b) Set Book

Study section **2.5.5** (pp. **124–129**) and complete Exercises 1, 2 and 3. If you are short of time, omit Exercise 2 and all but the last part of Exercise 3.

(c) Notes

(1) The geometrical interpretation of the general Taylor approximation is analogous to those of the tangent and quadratic approximations. (See Note c(2) in section 10.1.1 and Note c(1) in section 10.1.2.) If a polynomial is fitted to the function f at $(n + 1)$ equally spaced points

$$a, a + h, a + 2h, \ldots, a + nh,$$

and the limit as h approaches zero is taken, then the expression obtained for $f(x)$ is the Taylor approximation of degree n. Since the manipulation involved in this derivation is very tedious, you are advised not to attempt to derive this result for yourself.

(2) Note that the Taylor approximation of degree n is characterized by the equations

$$\begin{aligned} p(a) &= f(a)\\ p'(a) &= f'(a)\\ p''(a) &= f''(a)\\ &\ldots\\ p^{(n)}(a) &= f^{(n)}(a). \end{aligned}$$

The Taylor approximation of degree n satisfies these equations and, conversely, any polynomial of degree n which satisfies these equations *is* the Taylor approximation of degree n. (See Note c(2) in section 10.1.1, Note c(1) and Additional Exercise 2 in section 10.1.2.)

(3) It can be shown that

$$\lim_{x \to a} \frac{|f(x) - p(x)|}{|x - a|^k} = 0 \qquad (k = 1, 2, \ldots, n),$$

where $p(x)$ is the Taylor approximation of degree n. This follows from the $(n + 1)$ equations listed in Note (2) above.

(4) Exercise 3 and its solution (p. **129** and pp. **148–150**) brings out two important points. First, for some functions the Taylor approximation is exact if n is large enough. If

$$f : x \longmapsto (ax + b)^n \qquad (x \in R)$$

$(n \in Z^+)$, then the Taylor approximation of degree n about $x = 0$ is in fact the *binomial expansion* of $(ax + b)^n$. (See *Unit 2*, section 2.3.3.)

Second, for some functions the Taylor approximation about the point a is only reasonable within a limited neighbourhood of a. We must examine carefully the extent of the region in which we can use $p(x)$ as an approximation to $f(x)$. This point is taken up in sections 10.1.4 and 10.2.4.

(d) Self-Assessment Questions

1 Let f be the polynomial function

$$f : x \longmapsto x^3 - 5x^2 - 2x + 4 \qquad (x \in R).$$

Find values of the constants c_0, c_1, c_2, c_3 such that f may be specified in the form

$$f : x \longmapsto c_0 + c_1(x-2) + c_2(x-2)^2 + c_3(x-2)^3 \qquad (x \in R).$$

2 Derive an expression for the error in approximating the function

$$f : x \longmapsto \frac{1}{x} \qquad (x \in R^+)$$

over the interval $[a-\delta, a+\delta]$, $a > \delta > 0$, by a Taylor approximation of degree n.

Solutions

1 We have

$$\begin{aligned} f &: x \longmapsto x^3 - 5x^2 - 2x + 4 && (x \in R),\\ f' &: x \longmapsto 3x^2 - 10x - 2 && (x \in R),\\ f'' &: x \longmapsto 6x - 10 && (x \in R),\\ f^{(3)} &: x \longmapsto 6 && (x \in R)\\ f^{(n)} &: x \longmapsto 0 && (x \in R),\ n \geqslant 4. \end{aligned}$$

This means that the Taylor approximation of f about any point a in its domain terminates after four terms. So the required expression *is* the Taylor approximation of degree $n \geqslant 4$ about $x = 2$. Hence

$$\begin{aligned} c_0 &= f(2) = -12\\ c_1 &= f'(2) = -10\\ c_2 &= \tfrac{1}{2}f''(2) = 1\\ c_3 &= \tfrac{1}{6}f^{(3)}(2) = 1. \end{aligned}$$

2 The derived functions of f are

$$\begin{aligned} f' &: x \longmapsto -\frac{1}{x^2} && (x \in R^+),\\ f'' &: x \longmapsto \frac{2}{x^3} && (x \in R^+),\\ f^{(3)} &: x \longmapsto \frac{-3 \times 2}{x^4} && (x \in R^+)\\ f^{(4)} &: x \longmapsto \frac{4 \times 3 \times 2}{x^5} && (x \in R^+)\\ f^{(n)} &: x \longmapsto \frac{n!(-1)^n}{x^{n+1}} && (x \in R^+). \end{aligned}$$

Hence the Taylor approximation of degree n about a is

$$\begin{aligned} p_n(x) = {} & \frac{1}{a} - (x-a)\frac{1}{a^2} + \frac{1}{2}(x-a)^2 \frac{2}{a^3}\\ & - \frac{1}{3!}(x-a)^3 \frac{3 \times 2}{a^4} + \cdots \frac{+1}{n!}(-1)^n(x-a)^n \frac{n!}{a^{n+1}}. \end{aligned}$$

Let $x = a + \varepsilon$, where $\varepsilon \in [-\delta, \delta]$. Then the required error is

$$\begin{aligned} |f(x) - p_n(x)| &= \left| \frac{1}{a+\varepsilon} - \left(\frac{1}{a} - \frac{\varepsilon}{a^2} + \frac{\varepsilon^2}{a^3} - \cdots + (-1)^n \frac{\varepsilon^n}{a^{n+1}} \right) \right|\\ &= \frac{1}{(a+\varepsilon)a^{n+1}} \,|a^{n+1} - (a+\varepsilon)a^n + (a+\varepsilon)a^{n-1}\varepsilon\\ &\qquad\qquad + \cdots - (-1)^n(a+\varepsilon)\varepsilon^n|\\ &= \frac{|\varepsilon|^{n+1}}{(a+\varepsilon)a^{n+1}}. \end{aligned}$$

The maximum value of $|f^{(n+1)}(x)|$, $x \in [a-\delta, a+\delta]$, is

$$\left|\frac{(n+1)!(-1)^{n+1}}{(a-\delta)^{n+2}}\right| \geqslant \frac{(n+1)!}{(a+\varepsilon)a^{n+1}}.$$

Hence

$$\begin{aligned}\text{error} &= \frac{|x-a|^{n+1}}{(a+\varepsilon)a^{n+1}}\\ &\leqslant \frac{1}{(n+1)!}\,|x-a|^{n+1} \times \left(\begin{matrix}\text{maximum value of}\\ |f^{(n+1)}(x)| \text{ in } [a-\delta, a+\delta]\end{matrix}\right).\end{aligned}$$

We shall see in section 10.2.4 that this result, which we have obtained for a special case, holds for any function f with continuous derived functions of orders $1, \ldots, n+1$.

(e) Terminology

Defined in this section:

general Taylor approximation of degree n

(f) Additional Exercise (OPTIONAL MATERIAL)

Additional Exercise 3 (p. **144**) provides an opportunity to revise the ideas discussed in this section.

10.1.4 Error in the Tangent Approximation

(a) Purpose

To discuss an estimate of the error involved when a differentiable function f is approximated by the tangent approximation in the neighbourhood of a point a in its domain, and to state (without proof) Taylor's theorem for the tangent approximation.

(b) Notes

(1) In sections 10.1.1–10.1.3 we have defined the error in an approximation as the quantity

|exact value − approximation|.

In this section the *error* is defined as the *signed* quantity

approximation − exact value,

and its negative

exact value − approximation,

is called the *correction* to the approximation.

(2) From the self-assessment questions in section 10.1.1, we have already found the error in the tangent approximation for two simple functions, and we have found an expression which gives an upper bound on this error. The aim of section **2.5.6** is to extend this result to the general Taylor approximation.

(c) Set Book

Study section **2.5.6** (pp. **130–134**) and complete Exercises 1 and 2.

(d) Notes

(1) The result given on p. **131** (Inequality (1)) corresponds to the result which we established for two particular cases in the self-assessment questions of section 10.1.1. This is the statement of Taylor's theorem for the tangent approximation, namely that, if

$$|f''(t)| \leqslant B \qquad (t \in [a, x]),$$

then

$$|C_1(x)| = |f(x) - f(a) - (x-a)f'(a)| \leqslant \tfrac{1}{2}(x-a)^2 B.$$

We shall demonstrate the validity of Taylor's theorem for this particular case. (Taylor's theorem for the general *n*th degree polynomial approximation is discussed in section **2.5.7.**)

Let us consider the case $x > a$. We define the function

$$C_1 : t \longmapsto f(t) - [f(a) + (t-a)f'(a)] \qquad (t \in [a, x]).$$

The first and second derived functions of C_1 are

$$C_1' : t \longmapsto f'(t) - f'(a) \qquad (t \in [a, x]),$$

$$C_1'' : t \longmapsto f''(t) \qquad (t \in [a, x]).$$

Suppose that there exists a positive number B such that

$$|f''(t)| \leqslant B \qquad (t \in [a, x]).$$

Then, since $C_1''(t) = f''(t)$ and $C_1''(t) \leqslant |C_1''(t)|$, we have

$$C_1''(t) \leqslant B \qquad (t \in [a, x]).$$

We now integrate the above inequality from a to s, where $s \in [a, x]$. We obtain

$$\int_a^s t \longmapsto C_1''(t) \leqslant \int_a^s t \longmapsto B.$$

By the Fundamental Theorem of Calculus (*Unit 5*), it follows that

$$C_1'(s) - C_1'(a) \leqslant B(s-a),$$

since C_1' is a primitive of C_1''. But $C_1'(a) = 0$, so we have

$$C_1'(s) \leqslant B(s-a).$$

Since this inequality holds for all s in $[a, x]$, it follows that

$$C_1'(t) \leqslant B(t-a) \qquad (t \in [a, x]).$$

We now repeat the procedure and reduce the order of the derivative one further. Integrating this inequality from a to s, we obtain

$$\int_a^s t \longmapsto C_1'(t) \leqslant \int_a^s t \longmapsto B(t-a).$$

By the Fundamental Theorem of Calculus, we now have

$$C_1(s) - C_1(a) \leqslant \tfrac{1}{2}B(s-a)^2 \qquad (s \in [a, x]).$$

$C_1(a) = 0$, so we have

$$C_1(s) \leqslant \tfrac{1}{2}B(s-a)^2 \qquad (s \in [a, x]).$$

Similarly, starting with $-C_1''(t) \leqslant |C_1''(t)| \leqslant B$, we could obtain the inequality

$$-C_1(s) \leqslant \tfrac{1}{2}B(s-a)^2 \qquad (s \in [a, x]).$$

It follows that

$$|C_1(s)| \leqslant \tfrac{1}{2}B(s-a)^2 \qquad (s \in [a, x]).$$

A similar demonstration applies for $x < a$. The general Taylor theorem may be demonstrated similarly. (See section **2.5.7.**)

(2) The fact that we can place an upper bound on our error does *not* imply that the error is small. When we use the tangent approximation in a particular case, we must find the value of this upper bound in order to see if the error is acceptable. (See Example 1, p. **132.**)

Notice, however, that, since the error bound depends upon the quantity $(x-a)^2$, we can make the error as small as we please by restricting x to a region close to the point a.

(3) Note that our result gives an alternative proof of the statement

$$\lim_{x \to a} \frac{|f(x) - t(x)|}{|x - a|} = 0.$$

(See section 10.1.1, Note c(3).)

For

$$|f(x) - t(x)| = C_1(x)$$

and therefore

$$\lim_{x \to a} \frac{|f(x) - t(x)|}{|x - a|} \leqslant \lim_{x \to a} \frac{\frac{1}{2}B(x - a)^2}{|x - a|} = 0.$$

(e) Self-Assessment Question

Additional Exercise 4, section **2.5.9** (p. **144**) provides practice in using the Taylor theorem for linear polynomial approximations to estimate the error in a linear approximation.

(f) Terminology

Defined in this section:

error in an approximation
correction to an approximation
Taylor's theorem for linear polynomial approximations

10.2 OPTIONAL MATERIAL

10.2.1 The Convergence of an Iterative Method

(a) Purpose

To give a necessary condition for the convergence of the sequence generated using the recurrence formula

$$u_k = F(u_{k-1}) \qquad (k = 2, 3, 4, \ldots),$$

where u_1 is an approximate solution of the equation

$$x = F(x),$$

and to discuss the conditions under which we can use such an iterative process for solving equations of the form

$$f(x) = 0.$$

(b) Note

We started to consider the problem of solving equations of the form

$$f(x) = 0$$

in *Unit 1* (section 1.2.4, pp. 32–37) and *Unit 2* (section 2.2.4, pp. 32–35). If you did not study these sections, then work through them before attempting to complete the work below. For convenience, we shall start this section by summarizing the results obtained so far.

Given the equation $f(x) = 0$ with domain the real interval I, we define the function F by

$$F : x \longmapsto F(x) = f(x) + x \qquad (x \in I).$$

We saw in section 1.2.4 that the following conditions ensure that $F(x) = x$ has a solution in an interval $[a, b] \subseteq I$:

(i) F has domain and codomain $[a, b]$;
(ii) F is continuous on $[a, b]$.

In section 2.2.4 we considered how to select a form for $f(x) = 0$ so that the sequence generated by iteration using the corresponding function F converges to a solution of $f(x) = 0$. We found that, if F is a linear function, then the sequence converges provided that

$$|\text{the slope of the straight line graph of } F| < 1.$$

In the following material, we shall extend this result to a general function F. We shall then go on to consider the problem of selecting the interval $[a, b]$ so that it contains *only one* solution of the equation $f(x) = 0$.

(c) Set Book

Study section **2.5.2** (pp. **119–121**) and complete Exercise 1.

(d) Notes

(1) Notice the distinction between *necessary* conditions (conditions which *must* hold if $\underset{\sim}{u}$ does have a limit), and *sufficient* conditions (conditions which *ensure* the existence of a limit for the iterative sequence).

(2) Strictly speaking, the necessary condition produced takes it for granted that the function F is "well-behaved", that is, it is the kind of function to which it is reasonable to apply the tangent approximation and which satisfies the conditions imposed in Note (b) above. All these requirements are certainly met if we assume

that F is differentiable throughout its domain; however, this condition is stronger than we need, as we show below.

(3) In section 1.2.4 we examined two forms of functional iteration for solving the equation $x^2 - 2 = 0$. We can apply our test for convergence to these sequences.

In problem 3 (p. 34), we examined a functional iteration based upon

$$F : x \longmapsto x^2 - 2 + x \qquad (x \in [-2\tfrac{1}{3}, 1\tfrac{1}{3}]).$$

The corresponding derived function is

$$F' : x \longmapsto 2x + 1 \qquad (x \in [-2\tfrac{1}{3}, 1\tfrac{1}{3}]).$$

We have

$$|F'(-\sqrt{2})| = |2(-\sqrt{2}) + 1| > 1,$$

so the iterative sequence will not converge to the root at $-\sqrt{2}$ (and this is the only root in the interval $[-2\tfrac{1}{3}, 1\tfrac{1}{3}]$).

In Exercise 2 (p. 35), we examined a functional iteration based upon

$$F : x \longmapsto \frac{1}{x} + \frac{x}{2} \qquad (x \in [1, 1\tfrac{1}{2}]).$$

The corresponding derived function is

$$F' : x \longmapsto \frac{-1}{x^2} + \frac{1}{2} \qquad (x \in [1, 1\tfrac{1}{2}]).$$

We have

$$|F'(\sqrt{2})| = |-\tfrac{1}{2} + \tfrac{1}{2}| = 0 < 1,$$

so the iterative sequence will converge to the root at $\sqrt{2}$ (and this is the only root in the interval $[1, 1\tfrac{1}{2}]$).

(4) In section **2.5.2** we do not indicate how to obtain a *sufficient* condition for the convergence of the functional iteration. Our discussion in *Units 1* and *2* shows how to tackle this problem. We must find a condition which ensures that there is *only one* solution in the interval $[a, b]$. How can we find such a condition? What form will it take?

If F is differentiable on $[a, b]$, then one condition that will certainly meet our requirements is

$$|F'(x)| < 1 \text{ for all } x \in [a, b],$$

where $[a, b]$ is an interval on which F satisfies the conditions imposed in Note (b) above. This condition is certainly satisfied by a linear function with slope between -1 and $+1$ which we considered in *Unit 2*, section 2.2.4.

We can establish our assertion as follows.

Suppose $|F'(t)| < 1$ for all $t \in [a, b]$.
Then we may write $F'(t) \leqslant |F'(t)| \leqslant p$, where $p \in R$ and $0 < p < 1$.
We now integrate the above inequality from a to x, where $x \in [a, b]$. We obtain

$$\int_a^x t \longmapsto F'(t) \leqslant \int_a^x t \longmapsto p.$$

By the Fundamental Theorem of Calculus (*Unit 5*), it follows that

$$F(x) - F(a) \leqslant p(x - a).$$

Putting $x = u_{k-1}$, $F(u_{k-1}) = u_k$ and $F(a) = a$, we obtain

$$u_k - a \leqslant p(u_{k-1} - a).$$

Hence we obtain

$$u_k - a \leqslant p^{k-1}(u_1 - a),$$

and convergence is established, since

$$\lim_{k \text{ large}} (u_k - a) = 0$$

(because $\lim\limits_{k \text{ large}} p^k = 0$, see section 3.2.1).

How do we use this result? The approach is as follows.

Given an arbitrary equation $f(x) = 0$, sketch the graph of f so as to obtain some idea of the values of the solutions of the equation. Set up an appropriate function F so that it satisfies the following conditions:

(i) F is differentiable on $[a, b]$;
(ii) $|F'(x)| < 1$ for all $x \in [a, b]$;
(iii) $F(x) \in [a, b]$ for all $x \in [a, b]$.

The iterative sequence will then converge, as required.

10.2.2 The Newton-Raphson Process

(a) Purpose

To obtain an approximate method for solving the equation $f(x) = 0$.

(b) Set Book

Study section **2.5.3** (pp. **121–122**) and complete Exercise 1.

(c) Notes

(1) The method described is appealing on intuitive grounds for clearly if we are "sufficiently close" to a solution of $f(x) = 0$, then the graphical interpretation (p. **121**) suggests that the Newton-Raphson method will converge. (All the "steps" in the iteration are "trapped" between the curve, the tangent at the initial point, and the solution.) However, the derivation tells us nothing about the nature of the convergence, nor does it tell us how to find a starting value for the iteration which is sufficiently close to the solution we are seeking to ensure convergence to this solution. We consider this problem below.

(2) Another problem comes to light if we examine the form of the recurrence formula,

$$u_{k+1} = u_k - \frac{f(u_k)}{f'(u_k)}.$$

Clearly we shall encounter problems if $f'(u_k)$ is zero at one of the points u_k. When will this occur? We have seen in section **1.8.2** that $f'(u_k) = 0$ implies that the tangent to the graph of f is parallel to the x-axis at the point u_k. If we are already close to a root, then geometrically we can interpret $f'(u_k) = 0$ as implying that $f(x) = 0$ has two, or more, solutions close to the point $x = u_k$. By sketching appropriate curves for yourself, you should be able to see that, when the tangent to the graph of f is parallel, or nearly parallel, to the x-axis, then the linear approximation cannot be used to solve $f(x) = 0$, because it can lead to a value of x for which $|u_k - x|$ is very large.

In theory, the Newton-Raphson process can be used if u_k is close enough to a solution of $f(x) = 0$, and it will fail only if the equation has two or more coincident solutions. In practice, you are likely to encounter serious computational difficulties if the solutions are simply "close" rather than coincident.

(3) We now consider the problem of selecting a satisfactory initial approximation. The method to be used is the same as that used for the functional iteration problem. That is, we seek conditions on the function f that will enable us to identify an interval within which the initial approximation must lie. The labour

involved in establishing the appropriate conditions is considerable, and thus not appropriate to our introductory course. However, the intuitive meaning of the conditions are simple and straightforward, so we shall state the conditions to be satisfied, and give some explanation of their significance.

Let f be a function which is defined and has continuous derived functions, f', f'', on an interval $I \subseteq R$. This condition ensures that f is sufficiently "well-behaved" for us to be able to use the quadratic Taylor approximation to discuss the behaviour of the terms of the sequence which is generated by successive steps in the Newton-Raphson process.

Suppose that the following conditions are also satisfied on an interval $[a, b] \subseteq I$.

(i) $f(a)f(b) < 0$; that is, $f(a)$ and $f(b)$ have opposite signs.

(This condition ensures that there is at least one solution of $f(x) = 0$ in $[a, b]$.)

(ii) $f'(x) \neq 0$ for all $x \in [a, b]$.

(This condition, taken together with condition (i), ensures that there is only *one* solution of $f(x) = 0$ in $[a, b]$.)

(iii) Either

$$f''(x) \geqslant 0 \text{ for all } x \in [a, b]$$

or

$$f''(x) \leqslant 0 \text{ for all } x \in [a, b].$$

(This condition ensures that $f'(x)$ is either increasing or decreasing throughout $[a, b]$, which in turn enables us to show that the terms of the sequence $\underset{\sim}{u}$ are such that either $u_k \geqslant u_{k-1}$ or $u_k \leqslant u_{k-1}$, thus forming a steady progression of closer approximations to the limit.)

(iv) If c denotes the end-point of $[a, b]$ at which $|f'(x)|$ is smallest, then $\dfrac{|f(c)|}{|f'(c)|} \leqslant b - a$.

(This condition ensures that the tangent to the graph of f at the end-point intersects the x-axis within the interval $[a, b]$.)

Under this set of conditions, the Newton-Raphson method converges to the (only) solution of $f(x) = 0$ in $[a, b]$ for any choice of initial approximation $x_1 \in [a, b]$.

These conditions are sufficient but *not* necessary; convergence can be established for weaker sets of conditions. In practice, to use this result we must first sketch the graph of f and use it as a basis for selecting an interval which we can then show satisfies the conditions specified.

(4) One final point: although we derived the Newton-Raphson formula by a different technique to the basic functional iteration method (described in *Unit 1*, section 1.2.4), it does not represent a radically different idea. If you compare the solution of Exercise 1 of section **2.5.3** (p. **146**) with Exercise 2 of section 1.2.4 (p. 35), you see that the iteration formula which we established in both cases is identical. This observation underlines the fact that we can use many different methods to set up an iterative process to solve the equation $f(x) = 0$; for example, we would obtain yet another formula if we were to approximate the graph of f, near to a solution, by a chord rather than a tangent, and so on.

10.2.3 Power Series

(a) Purpose

To explain what is mean by a *power series* and the *sum* of a (convergent) power series, and to give the (convergent) Maclaurin expansion of some standard functions.

(b) Note

We introduced the concept of an infinite series and defined the meaning of the sum of a convergent infinite series in section 3.2.3 of *Unit 3*. We advise you to revise (or work through) that section before studying the material specified below.

(c) Set Book

Study section **2.5.8** (pp. **139–144**), and complete the exercises. (You completed Exercise 1 when you studied section 3.2.3 in *Unit 3*.) Omit Exercise 4 if you are short of time.

(d) Notes

(1) In section 3.2.3 of *Unit 3* we dealt mainly with infinite series that did not contain a variable. The exception was the geometric series

$$1 + r + r^2 + \cdots$$

We have seen that this series converges to the sum $1/(1-r)$ if $|r| < 1$ and that it diverges if $|r| \geqslant 1$: that is, the convergence depends on the value of the variable r.

In section **2.5.8** we discuss this more general type of infinite series in which a variable element is involved. The series

$$a_0 + a_1x + a_2x^2 + \ldots \qquad (x \in R)$$

is called a *power series*, because it is composed of powers of the variable x. When a fixed value is assigned to x, the power series reduces to the type of infinite series discussed in *Unit 3*. Thus a power series may be convergent for some values of x and divergent for other values of x. Since the sum of a convergent power series depends on x, it is denoted by $S(x)$ to show that it is a function of x.

(2) Exercise 2, and its solution (pp. **141** and **153–154**) illustrates the link between the Taylor (Maclaurin) approximation discussed in section 10.1.3 and the mathematics of infinite series.

If the function f has derivatives of *any* order, then the result of Exercise 2 tells us that, if the power series

$$f(a) + f'(a)(x-a) + \tfrac{1}{2}f''(a)(x-a)^2 + \cdots \qquad (x \in R)$$

has the property that $\lim\limits_{k \text{ large}} C_k(x) = 0$, then it converges to $f(x)$.

We call such a series the *Taylor series* (*expansion*) for the function f at the point a; if a is zero, then we call the series the *Maclaurin series* (*expansion*) for f. Some examples of Maclaurin series are given on page **143**. It can also be shown that, if the power series

$$b_0 + b_1(x-a) + b_2(x-a)^2 + \cdots$$

converges to $f(x)$, then the series *is* the Taylor series for f at the point a; that is

$$b_k = \frac{1}{k!}f^{(k)}(a).$$

(C.f. Note c(2) of section 10.1.1 and Note c(1) of section 10.1.2.)

(3) In section 3.2.5 of *Unit 3*, we discussed the problem of summing infinite series, and described some simple tests for determining whether a particular series is convergent or divergent. The techniques described there can be applied to the problem of summing power series. We see from Note (2) above, that we may

be able to obtain the sum of the series directly (by recognizing f) *if* we can establish that the series is absolutely convergent. We would like to be able to determine the values of x for which a given power series is absolutely convergent. For simplicity, we can consider a power series in the form

$$a_0 + a_1 x + a_2 x^2 + \cdots,$$

since the more general form

$$b_0 + b_1(x - a) + b_2(x - a)^2 + \cdots$$

can be regarded as a power series in the variable

$$X = x - a.$$

We can obtain a test for absolute convergence by making use of the ratio test described in section 3.2.5 (f) of *Unit 3*. We consider the ratio of two successive terms in the series. The result to be used can be stated as follows: If the series

$$c_1 + c_2 + c_3 + \cdots$$

is such that c_{n+1}/c_n is defined for all n and

$$\lim_{n \text{ large}} \left| \frac{c_{n+1}}{c_n} \right| = L,$$

then

if $L < 1$, the series is absolutely convergent;
if $L > 1$, the series is divergent;
if $L = 1$, the test gives no information about the convergence of the series.

If we apply the ratio test described above to the power series

$$a_0 + a_1 x + a_2 x^2 + \cdots,$$

then the condition for absolute convergence is

$$\lim_{n \text{ large}} \left| \frac{a_{n+1}}{a_n} \right| |x| < 1,$$

where $a_n \neq 0$.

If there exists a number $r > 0$ such that $\sum a_n x^n$ converges absolutely for $|x| < r$ and diverges for $|x| > r$, then r is called the *radius of convergence* of the power series, and the interval $]-r, r[$ is called the *interval of convergence*. Notice that the ratio test leaves us with an *open* interval, since the test is inconclusive when $|x| = r$. The behaviour of the series for $x = \pm r$ requires separate investigation. If the series is absolutely convergent for *all* values of x, then the radius of convergence is said to be *infinite*.

(4) In section **2.5.8**, we have used an alternative criterion to establish convergence, based on the result of Exercise 2 (pp. **141** and **153–154**); we examine the values of x for which

$$\lim_{k \text{ large}} C_k(x) = 0,$$

where $C_k(x)$ is the correction to an expansion. This test gives us the results summarized on page **143**. To apply this result, we often make use of the estimate for $C_k(x)$ derived in section **2.5.7**.

(e) Self-Assessment Questions

Investigate the convergence of the following power series:

1 $1 - \dfrac{x^2}{2} + \dfrac{x^4}{4} - \dfrac{x^6}{8} + \cdots$

2 $x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \dfrac{x^4}{4} + \cdots$

3 $1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$

4 $x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$

Solutions

1 The series is a geometric series with common ratio $-\frac{x^2}{2}$.

The series is absolutely convergent if

$$\left| -\frac{x^2}{2} \right| < 1 \qquad \text{i.e.} \qquad x \in]-\sqrt{2}, \sqrt{2}[.$$

If $x = \pm\sqrt{2}$, the series is divergent.

2 The nth term in the series is $(-1)^{n+1}x^n/n$. Using the ratio test, we obtain

$$\begin{aligned}\lim_{n\,\text{large}} \left| \frac{a_{n+1}}{a_n} \right| |x| &= \lim_{n\,\text{large}} \frac{n}{n+1} |x| \\ &= |x| \lim_{n\,\text{large}} \frac{n}{n+1} \\ &= |x|.\end{aligned}$$

Therefore the series is absolutely convergent for $|x| < 1$; that is, the radius of convergence is 1. For $|x| < 1$, the series is the Maclaurin series for the function

$$f: x \longmapsto \ln(1+x) \qquad (x \in]-1, 1[).$$

It can be shown that the series converges for $x = 1$ and diverges for $x = -1$.

3 The nth term in the series is $x^{n-1}/(n-1)!$
Using the ratio test, we obtain

$$\begin{aligned}\lim_{n\,\text{large}} \left| \frac{a_{n+1}}{a_n} \right| |x| &= \lim_{n\,\text{large}} \frac{(n-1)!}{n!} |x| \\ &= |x| \lim_{n\,\text{large}} \frac{1}{n} \\ &= |x| \times 0 \\ &= 0.\end{aligned}$$

Therefore the series is absolutely convergent for all values of x.

The series is the Maclaurin series for the exponential function.

4 The nth term in the series is $(-1)^{n+1}x^{2n-1}/(2n-1)!$ Using the ratio test, we obtain

$$\begin{aligned}\lim_{n\,\text{large}} \left| \frac{c_{n+1}}{c_n} \right| &= \lim_{n\,\text{large}} \frac{(2n-1)!}{(2n+1)!} |x|^2 \\ &= |x|^2 \lim_{n\,\text{large}} \frac{1}{2n(2n+1)} \\ &= |x|^2 \times 0.\end{aligned}$$

Therefore the series is absolutely convergent for all values of x.

The series is the Maclaurin series for the sine function.

(f) Terminology

Defined in this section:

power series
Taylor series (expansion)
Maclaurin series (expansion)
radius of convergence
interval of convergence

10.2.4 Taylor's Theorem

(a) Purpose

To demonstrate Taylor's theorem; that is, to show that the condition

$$|f^{(n+1)}(t)| \leqslant B \qquad (t \in [a, x])$$

is sufficient to imply that the correction associated with the Taylor expansion of a function f at the point a satisfies the inequality

$$|C_n(x)| \leqslant \frac{B}{(n+1)!}|x-a|^{n+1}.$$

(b) Set Book

Study section **2.5.7** (pp. **134–138**) and complete Exercise 1.

(c) Notes

(1) The method used to demonstrate the general Taylor theorem is a simple generalization of that used to demonstrate the Taylor theorem for linear approximations (section 10.1.4).

(2) Remember that the importance of this theorem is that it enables us to estimate the maximum error involved when we approximate a function by a Taylor (or Maclaurin) expansion.

(d) Self-Assessment Question

In the Maclaurin approximation of degree n for $f : x \longmapsto e^x$, how large does n have to be for the error in the approximation for e to be less than 0.001? (Assume that $e = 2.71828\ldots$)

Solution

e^x increases steadily as x increases, and $f^{(n)}(x) = e^x$ for all x. So in the interval [0, 1],

$$f^{(n+1)}(t) \leqslant e.$$

It follows that

$$\begin{aligned} \text{error} = C_n(1) &\leqslant \frac{e}{(n+1)!}(1-0)^{n+1} \\ &= \frac{2.71828}{(n+1)!} \end{aligned}$$

We find that, if we take the Maclaurin approximation of degree 6, and substitute $x = 1$, the error in the value we obtain for e is less than 0.001.

(e) Terminology

Defined in this section:

Taylor's theorem

10.2.5 Maxima and Minima of Functions of Two Real Variables

In this section, we return to the problem of the determination of local extrema of functions of two real variables, which we introduced in *Unit 7*, section 7.1.7. We use the Taylor series for a function of one real variable to construct a test which enables us to determine the *nature* of such extrema in certain cases.

So far in this unit we have written the Taylor series about $x = a$ in the form

$$f(x) = f(a) + (x - a)f'(a) + \tfrac{1}{2}(x - a)^2 f''(a) + \cdots,$$

that is, as an expression for $f(x)$, where x is "near" a *particular point a*. We now wish to use this result in the more general form

$$f(x + h) = f(x) + hf'(x) + \tfrac{1}{2}h^2 f''(x) + \cdots,$$

that is, as an expression for the image under f "near" the *general point x*.

Let F be a real function of the two real variables x and y. We shall assume that F has continuous partial derivatives of order 1, 2, 3, 4 with respect to x and y (alone or combined).

We define $F(a, b)$ to be a local *maximum* (*minimum*) value of $F(x, y)$ if

$$F(a + h, b + k) - F(a, b)$$

is *negative* (*positive*) for *all* small independent values of h and k.

We now use the Taylor series for a function of *one* real variable to investigate the sign of $F(x + h, y + k) - F(x, y)$. We do this by treating F as a function of each real variable in turn, *the other variable being held constant*. This means that the derivatives in the Taylor series will be *partial derivatives*. (See *Unit 7*, section 7.1.4.)

We shall adopt the following notation. We shall denote

$$F'_1(x, y) \text{ or } \frac{\partial F}{\partial x} \text{ by } F_x(x, y),$$

$$F'_2(x, y) \text{ or } \frac{\partial F}{\partial y} \text{ by } F_y(x, y),$$

$$F'_{11}(x, y) \text{ or } \frac{\partial^2 F}{\partial x^2} \text{ by } F_{xx}(x, y)$$

$$F'_{22}(x, y) \text{ or } \frac{\partial^2 F}{\partial y^2} \text{ by } F_{yy}(x, y),$$

$$F'_{12}(x, y) \text{ or } \frac{\partial^2 F}{\partial y\, \partial x} \text{ by } F_{yx}(x, y),$$

$$F'_{21}(x, y) \text{ or } \frac{\partial^2 F}{\partial x\, \partial y} \text{ by } F_{xy}(x, y),$$

and so on.

The conditions which we have assumed F to satisfy are sufficient to ensure that

$$F_{xy} = F_{yx}.$$

We begin by considering y and k as constants and treating F as a function of x alone. Using the Taylor series, we obtain

$$F(x + h, y + k) = F(x, y + k) + hF_x(x, y + k) + \tfrac{1}{2}h^2 F_{xx}(x, y + k) + \cdots \quad \text{(i)}$$

We now treat x as constant, and treat F as a function of y alone. Using the Taylor series, we obtain

$$F(x, y + k) = F(x, y) + kF_y(x, y) + \tfrac{1}{2}k^2 F_{yy}(x, y) + \cdots \quad \text{(ii)}$$

We now take equation (ii) and replace F by F_x to give

$$F_x(x, y + k) = F_x(x, y) + kF_{yx}(x, y) + \tfrac{1}{2}k^2 F_{yyx}(x, y) + \cdots \quad \text{(iii)}$$

Similarly, replacing F in (ii) by F_{xx}, we obtain

$$F_{xx}(x, y+k) = F_{xx}(x, y) + kF_{yxx}(x, y) + \tfrac{1}{2}k^2F_{yyxx}(x, y) + \cdots \qquad \text{(iv)}$$

Substituting equations (ii), (iii) and (iv) in equation (i), we obtain

$$\begin{aligned} F(x+h, y+k) = F(x, y) &+ kF_y(x, y) + hF_x(x, y) \\ &+ \tfrac{1}{2}k^2F_{yy}(x, y) + hkF_{yx}(x, y) \\ &+ \tfrac{1}{2}h^2F_{xx}(x, y) + \cdots \end{aligned}$$

Thus

$$\begin{aligned} F(x+h, y+k) - F(x, y) &= kF_y(x, y) + hF_x(x, y) \\ &\quad + \tfrac{1}{2}[k^2F_{yy}(x, y) + 2hkF_{yx}(x, y) + h^2F_{xx}(x, y)] + \cdots \end{aligned}$$

Now, if h and k are sufficiently small, the sign of

$$F(x+h, y+k) - F(x, y)$$

is the sign of

$$kF_y(x, y) + hF_x(x, y).$$

The last expression changes sign with the signs of h and k, so a necessary condition for (a, b) to be a local maximum or a local minimum is

$$kF_y(a, b) + hF_x(a, b) = 0 \text{ for } \textit{all } h, k;$$

that is,

$$F_x(a, b) = F_y(a, b) = 0.$$

When this condition is satisfied, the sign of

$$F(a+h, b+k) - F(a, b)$$

is the sign of

$$k^2F_{yy}(a, b) + 2hkF_{yx}(a, b) + h^2F_{xx}(a, b).$$

Thus, for a local maximum or minimum, this expression must have the same sign for all small values of h and k. Writing

$$\begin{aligned} A &= F_{yy}(a, b), \\ B &= F_{yx}(a, b), \\ C &= F_{xx}(a, b) \end{aligned}$$

and completing the square, we obtain

$$A\left[\left(k + \frac{B}{A}h\right)^2 + \left(\frac{AC - B^2}{A^2}\right)h^2\right], \ (A \neq 0).$$

The sign of this expression is the same as the sign of A, provided that $AC - B^2 \geqslant 0$ and h and k are not both zero.

Let us assume that $AC - B^2 > 0$, that is,

$$F_{yy}(a, b)F_{xx}(a, b) - (F_{yx}(a, b))^2 > 0.$$

Then

$$F(a+h, b+k) - F(a, b) > 0 \text{ if } F_{yy}(a, b) > 0$$

and

$$F(a+h, b+k) - F(a, b) < 0 \text{ if } F_{yy}(a, b) < 0.$$

Thus we have the following test:

If (a, b) is such that

$$F_x(a, b) = F_y(a, b) = 0$$

and

$$F_{yy}(a, b)F_{xx}(a, b) - (F_{yx}(a, b))^2 > 0,$$

then

(a, b) is a local minimum if $F_{yy}(a, b) > 0$,
(a, b) is a local maximum if $F_{yy}(a, b) < 0$.

Note that, by symmetry, the last two lines may be replaced by

(a, b) is a local minimum if $F_{xx}(a, b) > 0$,
(a, b) is a local maximum if $F_{xx}(a, b) < 0$.

This is the form of the test given in *Unit 7*, section 7.1.7.

10.3 SUMMARY

In this unit we have considered the Taylor approximation in some detail.

When you have completed work on the *core material*, you should be able to write down the general Taylor approximation to $f(x)$ about the point $x = a$ in the domain of f, for a suitable function f. You should also be able to give an upper bound on the error when the tangent approximation is used.

The *optional material* extends the work of earlier units.

ELEMENTARY MATHEMATICS FOR SCIENCE AND TECHNOLOGY

1 Sets, Mappings and Sequences
2 Functions and Limits
3 The Definite Integral
4 Differentiation
5 The Fundamental Theorem of Calculus
6 NO TEXT
7 Optimization Problems
8 Techniques and Applications of Integration
9 Operations and Relations
10 Taylor Approximation
11 Morphisms—Geometric Vectors
12 NO TEXT
13 Vector Spaces—Matrices
14 Differential Equations
15 Linear Equations
16 Complex Numbers
17 Second Order Differential Equations

035 01215 9